ISSN 1392-3730/eISSN 1822-3605

VILNIUS TECH
Vilnius Gediminas Technical University

JOURNAL of CIVIL ENGINEERING and MANAGEMENT

2025
Volume 30
Issue 1
Pages 38–51
https://doi.org/10.3846/jcem.2024.21886

A NOVEL HYBRID MODEL FOR PREDICTING THE BEARING CAPACITY OF PILES

Li TAO[1], Xinhua XUE[2]✉

[1]*School of Civil Engineering & Architecture, Wenzhou Polytechnic, Wenzhou 325000, P.R. China*
[2]*College of Water Resource and Hydropower, Sichuan University, Chengdu 610065, P.R. China*

Article History:
- received 31 October 2023
- accepted 14 June 2024
- first published online 22 October 2024

Abstract. Due to the uncertainty of soil condition and pile design characteristics, it is always a challenge for geotechnical engineers to accurately determine the bearing capacity of piles. The main objective of this study is to propose a hybrid model coupling least squares support vector machine (LSSVM) with an improved particle swarm optimization (IPSO) algorithm for the prediction of bearing capacity of piles. The improved PSO algorithm was used to optimize the LSSVM hyperparameters. The performance of the IPSO-LSSVM model was compared with seven artificial intelligence models, namely adaptive neuro-fuzzy inference system (ANFIS), M5 model tree (M5MT), multivariate adaptive regression splines (MARS), gene expression programming (GEP), random forest (RF), regression tree (RT) and a stacked ensemble model. Six statistical indices (e.g., coefficient of determination (R^2), mean absolute error (MAE), root mean squared error (RMSE), relative root mean squared error (RRMSE), BIAS and discrepancy ratio (DR)) were used to evaluate the performance of the models. The R^2, MAE, RMSE, RRMSE and BIAS values of the IPSO-LSSVM model were 1, 4.27 kN, 6.164 kN, 0.005 and 0, respectively, for the training datasets and 0.9977, 22 kN, 36.03 kN, 0.0275 and –11, respectively, for the testing datasets. Compared with the ANFIS, MARS, GEP, M5MT, RF, RT and the stacked ensemble models, the proposed IPSO-LSSVM model shows high accuracy and robustness on the test datasets. In addition, the sensitivity, uncertainty, reliability and resilience of the IPSO-LSSVM model were also analyzed in this study.

Keywords: bearing capacity, pile, adaptive neuro-fuzzy inference system, least squares support vector machine, stacked ensemble model.

✉Corresponding author. E-mail: *13708026185@163.com*

1. Introduction

In geotechnical engineering, when the soil condition is not enough to support the design load, it is necessary to use pile to improve the bearing capacity of the foundation (Momeni et al., 2014). However, the traditional pile bearing capacity test has many disadvantages such as time consuming and high cost. In addition, static load tests are difficult to perform in some cases (Teh et al., 1997). To accurately predict the bearing capacity of piles, many empirical models have been developed in recent years. For example, Pal and Deswal (2010) studied the bearing capacity of piles by using Gaussian process regression model. Sheil and McCabe (2016) conducted a detailed study of the behavior of individual pile and pile groups in clay. Li et al. (2017) studied the time-dependent bearing capacity of a jacked pile. Rezazadeh and Eslami (2017) proposed an empirical model for determining shaft bearing capacity of semi-deep foundations socketed in rocks. Salgado et al. (2017) used an improved soil reaction model to predict a one-dimensional pile driving simulation and obtained a reliable pile driving formula. Luo et al. (2018) proposed a practical analysis method for the pile raft foundation. Wang et al. (2018) investigated the bearing capacity of large diameter monopiles in sandy soil by conducting a series of centrifuge tests.

In recent years, with the development of artificial intelligence (AI), machine learning techniques such as artificial neural networks (ANNs), fuzzy logic, adaptive neuro-fuzzy inference system (ANFIS), support vector machine (SVM), least squares support vector machine (LSSVM), M5 model tree (M5MT), multilayer perceptron (MLP), multivariate adaptive regression splines (MARS), random forest (RF), regression tree (RT), etc., and their combinations with evolutionary algorithms (e.g., genetic algorithms (GAs), genetic programming (GP), evolution strategies, evolutionary programming, ant colony optimization, particle swarm optimization (PSO), etc.), have been widely used in many

files such as viscosity prediction of TiO_2/water nanofluid (Ahmadi et al., 2020), estimation of densities and viscosities of amine-based solutions (Haratipour et al., 2017), natural gas water content estimation (Baghban et al., 2016a), prediction of true vapor pressure of petroleum products (Baghban et al., 2016b), modeling of viscosity for mixtures of Athabasca bitumen and heavy n-alkane (Baghban et al., 2016c), estimation of CO_2-Brine diffusivity (Bemani et al., 2020a), prediction of sulfur solubility in supercritical sour gases (Bemani et al., 2020b), precipitation estimation of asphaltene during different production processes (Baghban & Khoshkharam, 2016), estimation of acid solvent solubility in supercritical CO_2 (Bemani et al., 2020c), prediction of energy conservation benefits in excess air controlled gas-fired systems (Bahadori et al., 2016), determination of CO_2 absorption in polyionic liquids (Kardani et al., 2018), pile bearing capacity prediction (Das & Basudhar, 2006; Lin et al., 2009; Kalinli et al., 2011; Baziar et al., 2012; Alkroosh & Nikraz, 2012, 2014; Armaghani et al., 2017; Nejad & Jaksa, 2017; Murlidhar et al., 2020; Pham et al., 2020; Benbouras et al., 2021; Yong et al., 2021; Amjad et al., 2022), scour depth around bridge piers (Najafzadeh & Barani, 2011; Najafzadeh et al., 2013; Najafzadeh & Azamathulla, 2013a), scour studies of pile groups (Najafzadeh & Azamathulla, 2013b; Najafzadeh, 2015; Homaei & Najafzadeh, 2020; Najafzadeh & Oliveto, 2021), estimation of energy efficiency of flow-dissipating vortex dropshaft (Najafzadeh & Mahmoudi-Rad, 2024).

For example, Momeni et al. (2014) predicted the pile bearing capacity using a hybrid GA-based ANN model. Das and Basudhar (2006) investigated the undrained lateral load capacity of piles in clay using an ANN model. Lin et al. (2009) investigated the failure potential of highway slopes using an ANN model. Kalinli et al. (2011) used ANN and ant colony optimization to study the ultimate bearing capacity of shallow foundations. Baziar et al. (2012) predicted the pile shaft resistance by using ANN and nonlinear multi regression models. Alkroosh and Nikraz (2012) studied the axial capacity of driven piles in cohesive soils by using intelligent computing methods. Alkroosh and Nikraz (2014) developed a new approach for predicting the dynamic capacity of piles. Armaghani et al. (2017) used a hybrid PSO-ANN model to predict the ultimate bearing capacity of piles. Nejad and Jaksa (2017) studied the load-settlement behavior of single piles using ANNs models and CPT data. Pham et al. (2020) used ANN and RF algorithm to investigate the ultimate axial bearing capacity of driven piles. Yong et al. (2021) developed three soft computing techniques, including ANFIS, GP tree-based, and SA-GP to predict the ultimate bearing capacity of piles. Amjad et al. (2022) predicted the bearing capacity of piles using an extreme gradient boosting algorithm. Murlidhar et al. (2020) investigated the effects of PSO and GA on ANN results in predicting pile bearing capacity. Benbouras et al. (2021) studied the bearing capacity of the driven pile by using advanced machine learning techniques.

In light of the above-mentioned state-of-the-arts, the broader scientific literature, as well as the author's understanding, little research so far has dealt with the hybrid model which combines an improved PSO algorithm and LSSVM for the prediction of pile bearing capacity. Therefore, this study presents a novel hybrid IPSO-LSSVM model to predict the bearing capacity of piles. The improved PSO algorithm was used to optimize the LSSVM hyperparameters. The performance of the IPSO-LSSVM model was compared with seven AI models, namely ANFIS, M5MT, MARS, gene expression programming (GEP), RF, RT and a stacked ensemble model. Six statistical indices were used to evaluate the performance of the models. In addition, the sensitivity, uncertainty, reliability and resilience of the hybrid IPSO-LSSVM model were also analyzed in this study.

2. Methodology

2.1. Least squares support vector machine (LSSVM)

Assume a given data set $\left[x_i, y_i\right]\left(i=1,2,...,N\right)$, where x_i is the input data, y_i is the output data, and N denotes the number of training samples. The learning objective of LSSVM can be transformed into the following optimization problem (Suykens et al., 2001):

$$\text{Minimize} \qquad J\left(w,\varsigma\right)=\frac{1}{2}w^{\mathrm{T}}w+\frac{1}{2}\lambda\sum_{i=1}^{N}\varsigma_i^{\,2} \tag{1}$$

$$\text{Subjected to} \quad y_i=w^{\mathrm{T}}\varphi\left(x_i\right)+b+\varsigma_i,\ \ i=1,...,N, \tag{2}$$

where w is the weight matrix, λ is a constant of regularization, ς_i is a positive slack variable, b is the bias, $\varphi(x)$ is the nonlinear transformation that maps input data x_i to high dimensional feature space.

According to Mercer's condition (Mercer, 1909), the mapping function $\varphi(x)$ can be characterized by a kernel function $K\left(x_i,x_j\right)=\varphi\left(x_i\right)^{\mathrm{T}}\varphi\left(x_j\right)$. Therefore, the LSSVM model can be written as:

$$y\left(x\right)=\sum_{i=1}^{N}\alpha_i K\left(x,x_i\right)+b. \tag{3}$$

The radial basis function (RBF) kernel is often used:

$$K\left(x,x_i\right)=\exp\left(-\frac{\left\|x_i-x_j\right\|^2}{2\sigma^2}\right), \tag{4}$$

where α_i is the multiplier; σ is the RBF kernel parameter.

2.2. Improved particle swarm optimization (IPSO)

Kennedy and Eberhart first proposed the PSO algorithm in 1995. In PSO algorithm, the update of the position (x_{ij}) and velocity (v_{ij}) of every particle is carried out by the following equations (Kennedy & Eberhart, 1995):

$$x_{ij}^{t}=x_{ij}^{t-1}+v_{ij}^{t}; \tag{5}$$

$$v_{ij}^{t}=\kappa v_{ij}^{t-1}+c_1r_1\left(p_{ij}-x_{ij}^{t-1}\right)+c_2r_2\left(g_j-x_{ij}^{t-1}\right), \tag{6}$$

where r_1 and r_2 denote two random numbers; c_1 and c_2 denote acceleration coefficients; p_{ij} and g_j denote the best location found by every particle and the whole swarm, respectively. κ is the factor of inertia weight. Herein, an improved κ was proposed and written as:

$$\kappa = \left(\kappa_{end} - \kappa_{start}\right)\left(\frac{n_{\max} - n_i}{n_i}\right) + \kappa_{start}, \quad (7)$$

where κ_{end} is the final inertia weight; κ_{start} is the initial inertia weight; n_i isthe ith iteration number; $n_{\max}$ is the maximum iterations.

2.3. IPSO-LSSVM

In this study, the improved PSO algorithm was used to optimize the regularization parameter λ and kernel parameter σ, which play an important role in LSSVM. The main steps of the hybrid IPSO-LSSVM algorithm can be briefly summarized as follows.

- Step 1: Initialize the population of particles with random position and velocities.
- Step 2: Selection of fitness function. Herein, the root mean squared error (RMSE) is taken as the fitness function.
- Step 3: Update the velocity and position of each particle according to Eqns (5)–(7).
- Step 4: If the stop criteria is satisfied, the procedure of calculation ends; otherwise, return to Step 3.

The flowchart of the hybrid IPSO-LSSVM model is shown in Figure 1.

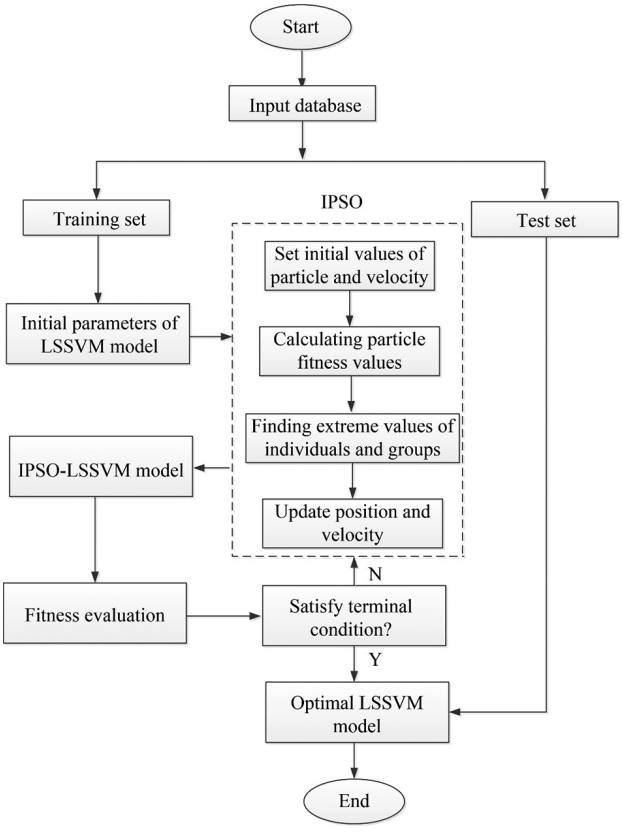

Figure 1. Flowchart of the IPSO-LSSVM algorithm

2.4. Adaptive neuro-fuzzy inference system (ANFIS)

The architecture of an ANFIS (as shown in Figure 2) is generally composed of premise and conclusion parts: (i) the fuzzy layer, (ii) the rule layer, (iii) the normalization layer, and (iv) the defuzzification layer.

The expressions of the five layers can be written as follows (Chiu, 1994).

Layer 1:

$$O_1^i = \mu_{A_j^k}\left(x_j\right), \quad j = 1,2,...,m; \quad k = 1,2,...,n; \quad i = n \times m, \quad (8)$$

where x_j are the inputs to node j, A_j^k are the linguistic labels characterized by membership functions (MFs) $\mu_{A_j^k}$.

Layer 2:

$$O_2^j = w_i = \Pi\mu_{A_j^i}\left(x_j\right), \quad j = 1,2,...,m; \quad i = 1,2,...,n, \quad (9)$$

where w_i denotes the firing strength of a rule.

Layer 3:

$$O_3^i = \bar{w}_i = \frac{w_i}{\sum_{i=1}^{n} w_i}, \quad (10)$$

where $\bar{w}_i$ denotes the firing strengths of normalization.

Layer 4:

$$O_4^i = \bar{w}_i f_i = \bar{w}_i\left(p_0^i + p_1^i x_1 + p_2^i x_2 + ... + p_m^i x_m\right), \quad i = 1,2,...,n, \quad (11)$$

where $\left\{p_0^i, p_1^i, ..., p_m^i\right\}$ are often referred to as consequent parameters.

Layer 5:

$$O_5^i = \sum_{i=1}^{n} \bar{w}_i f_i = \frac{\sum_{i=1}^{n} w_i f_i}{\sum_{i=1}^{n} w_i}. \quad (12)$$

2.5. M5 model tree (M5MT)

The M5MT technique, originally developed by Quinlan in 1992, is a robust approach to modeling and predicting phenomena that deals with complex implicit prob-

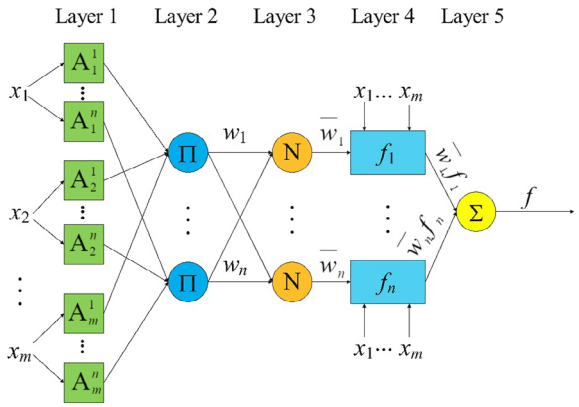

Figure 2. Structure of ANFIS

lems by dividing them into smaller ones (Najafzadeh & Oliveto, 2021; Homaei & Najafzadeh, 2022). In M5 model tree, there are two stages: (i) the input space is divided into sub regions, and (ii) the tree is built using the data from each sub region. First, the input space is divided into sub regions, and the data of each sub region is linearly fitted. Based on the input data points of the sub region, the regression tree is constructed in the second stage. In this tree, the leaves are at the bottom and the roots are at the top.

Reducing standard deviation (*RSD*) is the criterion in the splitting stage of the M5MT, and the expression of *RSD* can be written as follows (Quinlan, 1992):

$$RSD = sd(X) - \sum \frac{|X_i|}{|X|} sd(X_i), \tag{13}$$

where X is the set of samples which enter the node, X_i shows the examples' subset corresponding to ith value of the potential set and *sd* is the standard deviation. In the first stage, the M5MT selects the partition with the largest *RSD* after checking all available partitions. However, after this process, a larger tree-like structure is obtained. In the second stage, the overgrown trees obtained in the first stage are pruned and subtrees are used to replace the linear regression models. Further descriptions of M5MT can be found in Quinlan (1992).

2.6. Multivariate adaptive regression splines (MARS)

As a nonlinear nonparametric regression method, MARS models the nonlinear relationships between different variables through a series of piecewise linear splines of different gradients. The details of MARS can be summarized as follows (Friedman, 1991).

Considering the general model form of the relationship between the input x and the target output y as follows (Friedman, 1991):

$$y = f(x) = a_0 + \sum_{m=1}^{M} a_m B_m(x), \tag{14}$$

where a_0 is a constant; a_m is the coefficient of basis functions (BFs) B_m.

MARS uses the following bilateral truncated power functions as the spline BFs (Friedman, 1991):

$$b_q^{\pm}(x-t) = \left[\pm(x-t)\right]_+^q, \tag{15}$$

where t is the knot location, q is the order of the spline.

The analysis of variance decomposition of the MARS model is given by the following equation (Friedman, 1991):

$$f(x) = a_0 + \sum_{K_m=1} f_i(x_i) + \sum_{K_m=2} f_{ij}(x_i, x_j) + \sum_{K_m=3} f_{ijk}(x_i, x_j, x_k) + \ldots . \tag{16}$$

The first sum $\sum_{K_m=1} f_i(x_i)$ is over all BFs that involve only a single variable. The second sum $\sum_{K_m=2} f_{ij}(x_i, x_j)$ is over all BFs that involve exactly two variables. Similarly, the third sum $\sum_{K_m=3} f_{ijk}(x_i, x_j, x_k)$ represents (if present) the contributions from three variable interactions.

Equation (16) involves the sum of the product of functions, and the form is similar to the following (Friedman, 1991):

$$b(x|s,t) = \left[s(x-t)\right]_+. \tag{17}$$

To solve this problem, the strategy is to replace each such function with a corresponding truncated cubic function of the form (Friedman, 1991):

$$BF(x \mid s=+1, t_-, t, t_+) = \begin{cases} 0 & x \le t_-, \\ p_+(x-t_-)^2 + r_+(x-t_-)^3 & t_- < x < t_+, \\ x-t & x \ge t_+, \end{cases}$$

$$BF(x \mid s=-1, t_-, t, t_+) = \begin{cases} -(x-t) & x \le t_-, \\ p_-(x-t_+)^2 + r_-(x-t_+)^3 & t_- < x < t_+, \\ 0 & x \ge t_+ \end{cases} \tag{18}$$

with $t_- < x < t_+$.

Setting

$$p_+ = (2t_+ + t_- - 3t)/(t_+ - t_-)^2,$$
$$r_+ = (2t - t_+ - t_-)/(t_+ - t_-)^3,$$
$$p_- = (3t - 2t_- - t_+)/(t_- - t_+)^2,$$
$$r_- = (t_- + t_+ - 2t)/(t_- - t_+)^3 \tag{19}$$

causes $BF(x \mid s, t_-, t, t_+)$ to be continuous and have continuous first derivatives. More details of MARS can be found in Friedman (1991).

2.7. Gene expression programming (GEP)

The GEP model is a new form of genetic programming (GP). In contrast to the GP model, the individuals of the GEP model are designed to have a linear string (genome or chromosome) of a certain length. According to Ferreira (2001), chromosomes contain one or more genes, and a gene consists of a head and a tail. A function or terminal is a GEP symbol. Both the functions and terminals are stored in the head of the genes, while the tail contains only the terminals. More details of GEP can be found in Ferreira (2001).

2.8. Random forest (RF)

RF refers to a classifier that uses multiple trees to train and predict samples. RF is a supervised learning method that builds decision trees on data samples, generates and accumulates predictions for each sample, and then votes on

the best of them. When developing an RF model, two hyperparameters need to be determined, namely the depth of the decision tree and the number of decision trees. The RF algorithm is briefly described as follows (Daneshvar & Behnood, 2020):

(1) The bagging method is used to randomly extract samples from the original training set and construct *k* sample subsets.

(2) Using the random subspace method, *F* features are randomly extracted from all *X* feature attributes for node splitting, and the regression decision tree is constructed.

(3) Based on the concept of ensemble learning, the output of each regression decision tree is averaged to obtain the prediction results, as shown in Eqn (20) (Daneshvar & Behnood, 2020):

$$\bar{h}_{\mathrm{RF}}(x)=\frac{1}{T}\sum_{t=1}^{T}\left\{h(\theta_t,x)\right\}, \tag{20}$$

where $\bar{h}_{\mathrm{RF}}(x)$ represents the predicted result, T represents the number of regression decision trees, $h(\theta_t,x)$ represents the output based on θ_t and x, θ_t represents an independent and identically distributed random vector, and x represents an independent variable.

2.9. Regression tree (RT)

RT has been widely used in machine learning and data mining community. Given a target data for prediction, RT is constructed by recursively partitioning a data set and fitting a simple model to each partition. As a result, the partitioning can be represented graphically as a decision tree. Regression trees are for dependent variables that take continuous or ordered discrete values, with prediction error typically measured by the squared difference between the observed and predicted values. In practice, the performance of RT during the tree building/prediction phase relies heavily on the local mean of a single node sample while ignoring global information from different nodes, which also plays an important role. More details of RT can be found in Breiman et al. (1984).

2.10. Stacked ensemble model

Stacking is an ensemble learning method that combines multiple classification or regression models through a meta-classifier or a meta-regressor. Specifically, individual models were first developed on the training datasets and then, based on the results of each model, a meta-learner was used to develop the stacked ensemble model. In this study, seven AI models including IPSO-LSSVM, ANFIS, M5MT, MARS, GEP, RF and RT were first considered to develop the models for pile bearing capacity prediction. Then, according to the predicted results, three models with good performance are selected from the seven models to form a stacked ensemble model for more robust comparisons. More details of the proposed stacked ensemble model can be found in Section 4.

2.11. Statistical indices

In this study, six statistical indices, namely coefficient of determination (R^2), root mean squared error (RMSE), mean absolute error (MAE), relative root mean squared error (RRMSE), BIAS and discrepancy ratio (DR) (Najafzadeh et al., 2016; Najafzadeh & Oliveto, 2021; Homaei & Najafzadeh, 2022) were used to evaluate the prediction performance of the models. The expressions of these six statistical indices can be written as follows (Najafzadeh et al., 2016; Najafzadeh & Oliveto, 2021; Homaei & Najafzadeh, 2022):

$$R^2=1-\frac{\sum_{i=1}^{n}\left(y_{obs,i}-y_{pred,i}\right)^2}{\sum_{i=1}^{n}\left(y_{obs,i}-\bar{y}_{obs}\right)^2}; \tag{21}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}\left(y_{pred,i}-y_{obs,i}\right)^2\right)}; \tag{22}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|y_{pred,i}-y_{obs,i}\right|; \tag{23}$$

$$\mathrm{RMMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{pred,i}-y_{obs,i})^2}\ /\ \bar{y}_{obs}; \tag{24}$$

$$\mathrm{BIAS}=\bar{y}_{pred}-\bar{y}_{obs}; \tag{25}$$

$$\mathrm{DR}=\log\left(\frac{y_{pred}}{y_{obs}}\right), \tag{26}$$

where $y_{pred,i}$ and $y_{obs,i}$ represent the predicted and observed results, respectively. $\bar{y}_{pred}$ and $\bar{y}_{obs}$ represent the average value of the predicted and observed results, respectively. n represents the total sample size.

3. Database

In this study, the following four parameters of piles, i.e., pile diameter (D), embedded length of pile (L), the average lateral resistance of pile ($\bar{q}_{sik}$) and ultimate end bearing resistance of pile (q_{pk}) were considered as the inputs, while the pile bearing capacity (R_{uk}) was the output. The database collected by Zheng et al. (2006) was used to verify the model, and the details of which are summarized in Appendix Table A1.

Table 1. Statistical results of the experimental data

Parameter	Mean	Max	Min	Standard deviation
D/mm	485	600	400	66
L/m	39	59	24	10
$\bar{q}_{sik}$/kPa	15	21.1	8.4	4
q_{pk}/kPa	1704	3400	550	824
R_{uk}/kN	1332	2700	450	783

Table 1 lists the statistical results of the collected experimental data. The Pearson correlation coefficients between different parameters are shown in Table 2. As observed from Table 2, there is a significant correlation between ultimate bearing capacity and four input variables, and their correlation coefficients are all greater than 0.7.

Table 2. Pearson correlation coefficients between different parameters

Parameter	D/mm	L/m	$\overline{q}_{sik}$/kPa	q_{pk}/kPa	R_{uk}/kN
D/mm	1				
L/m	0.632	1			
$\overline{q}_{sik}$/kPa	0.698	0.671	1		
q_{pk}/kPa	0.518	0.452	0.547	1	
R_{uk}/kN	0.855	0.805	0.907	0.720	1

4. Results and discussion

4.1. Development of models

In this study, the optimal parameters of the IPSO-LSSVM model are as follows: acceleration coefficients $c_1 = c_2 = 2$, swarm size $N_p = 10$, maximum iterations $n_{max} = 100$, and the RBF kernel parameter $\sigma = 0.7448$. The optimal parameter of ANFIS and GEP models are shown in Tables 3 and 4, respectively.

Table 3. Different parameter types and their values used for training ANFIS

ANFIS parameter type	Setting
MF type	Gaussian
Number of fuzzy rules	4
Number of linear parameters	4
Number of nonlinear parameters	35
Number of training data pairs	50
Number of testing data pairs	30
Number of nodes	41
Total number of parameters	39

Table 4. Optimal parameters of the GEP model

Parameter	Setting	Parameter	Setting
Population size	1000	Chromosome length	50
Linking function	addition (+)	Number of genes	20
Gene transposition rate	0.7	IS transposition rate	0.3
Gene recombination rate	0.7	RIS transposition rate	0.3
Two-point recombination rate	0.4	Head size	15
One-point recombination rate	0.4	Mutation rate	0.1

4.2. Performance comparison of seven AI models

This section presents the performance comparison results of seven AI models. The forecasting performance comparisons of these seven AI models are displayed in Figure 3, Tables 5 and 6, respectively. Figure 4 to Figure 6 plot the Violin diagram, Taylor diagram and DR diagram of these seven AI models, respectively.

Table 5. Performance comparison of seven AI models on the training set

Model	MAE/kN	RMSE/kN	RRMSE	R^2	BIAS
IPSO-LSSVM	4.27	6.164	0.005	1	0
MARS	20.76	28.713	0.021	0.999	0
GEP	195.30	285.77	0.213	0.773	–61
M5MT	21.20	33.374	0.025	0.998	0
RF	0	0	0	1	0
RT	0	0	0	1	0
ANFIS	0.03	0.064	0	1	0

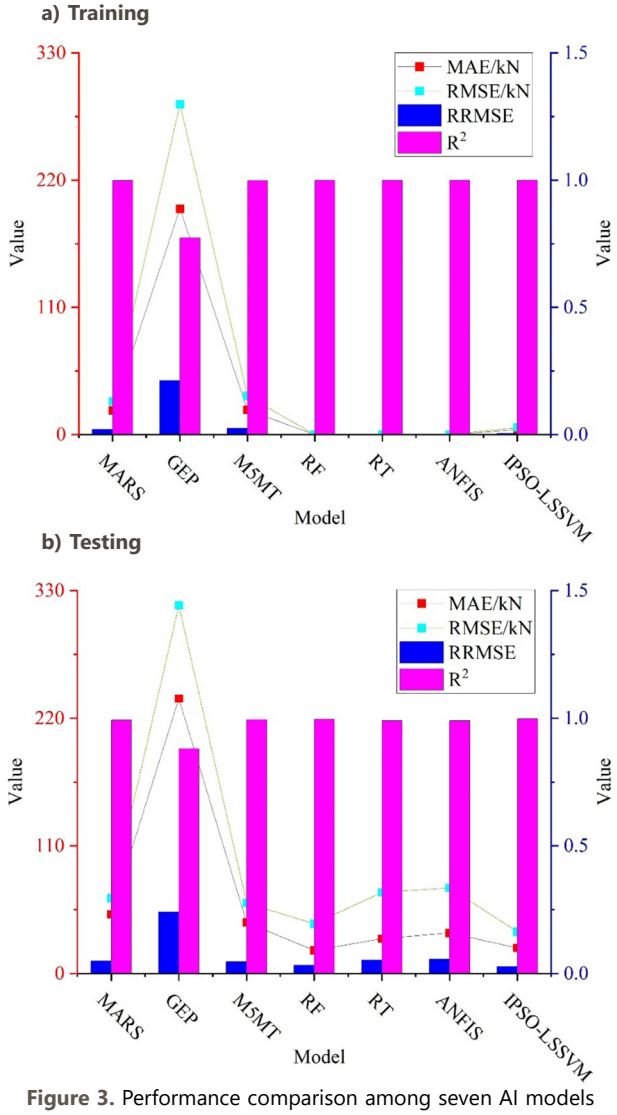

Figure 3. Performance comparison among seven AI models

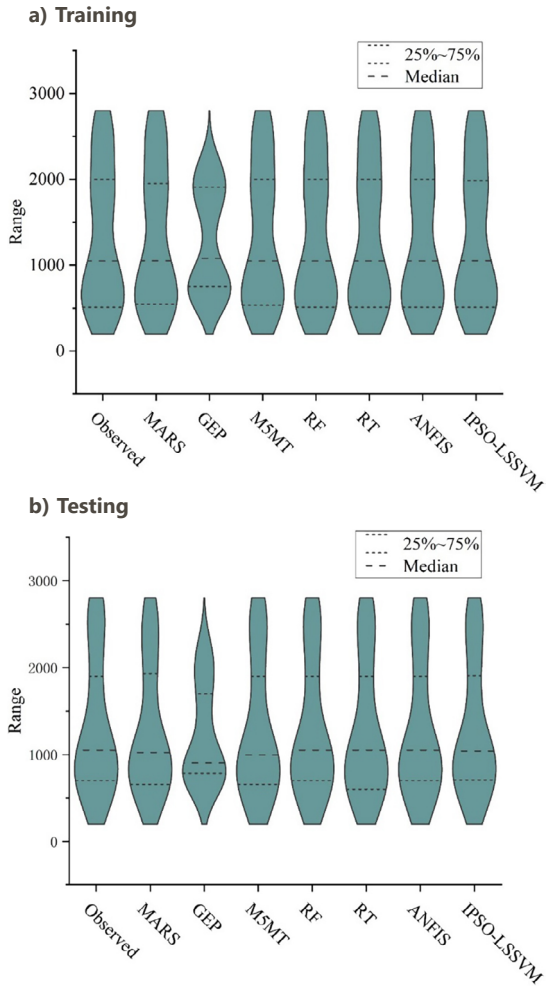

Figure 4. Violin diagram of seven AI models

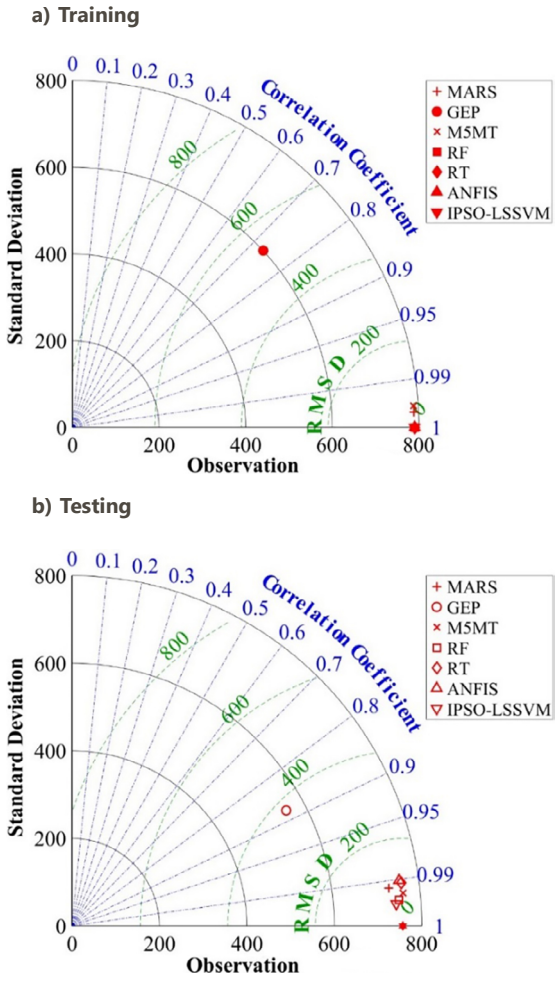

Figure 5. Taylor diagram of seven AI models

Table 6. Performance comparison of seven AI models on the test set

Model	MAE/kN	RMSE/kN	RRMSE	R^2	BIAS
IPSO-LSSVM	22	36.03	0.0275	0.9977	–11
MARS	51	64.839	0.0494	0.9930	–16
GEP	237	317.342	0.2419	0.8807	–94
M5MT	44	60.752	0.0463	0.9951	–29
RF	20	42.817	0.0326	0.9968	–13
RT	30	70	0.0534	0.9915	–25
ANFIS	35	73.883	0.0563	0.9905	–24

As observed from Figure 3, Tables 5 and 6, MARS, M5MT, RT, RF, ANFIS and IPSO-LSSVM models all have good effects in the training stage, where RT model (R^2 = 1, RMSE = 0 kN, MAE = 0 kN, RRMSE = 0, BIAS = 0) and RF model (R^2 = 1, RMSE = 0 kN, MAE = 0 kN, RRMSE = 0, BIAS = 0) had the best effect, followed by IPSO-LSSVM model (R^2 = 1, RMSE = 6.164 kN, MAE = 4.27 kN, RRMSE = 0.005, BIAS = 0). However, the training effect difference between IPSO-LSSVM and RT and RF models is very small and can be considered to be almost the same. In the test stage, IPSO-LSSVM model (R^2 = 0.9977, RMSE = 36.03 kN, MAE = 22kN, RRMSE = 0.0275, BIAS = –11) has the best effect and the closest distance to the observed value, which indicates that although the training effect of IPSO-LSSVM is slightly inferior to that of RF and RT models, its prediction ability is the most outstanding.

In addition, it can be seen from Figure 4 that in the training stage and the test stage, the data structure of the GEP model and the observed value is quite different, while the data structure of MARS, M5MT, RT, RF, ANFIS and IPSO-LSSVM models and the observed value are the same, indicating that these models have better prediction effect. As can be seen from Figure 5, the closer the Taylor chart model is to the observed value, the better the model effect is. It is found that in the training stage and the test stage, the GEP model has the farthest distance from the observed value and the worst effect. As can be seen from Figure 6, the closer DR is to 0, the better the prediction effect is. It can be found that in the training stage, the DR values of MARS, M5MT, RF, RT, ANFIS and IPSO-LSSVM models are all closely around 0, among which RT and RF are the best. However, in the test phase, the DR values of IPSO-LSSVM model are closest to 0, indicating that the IPSO-LSSVM model has the best prediction effect.

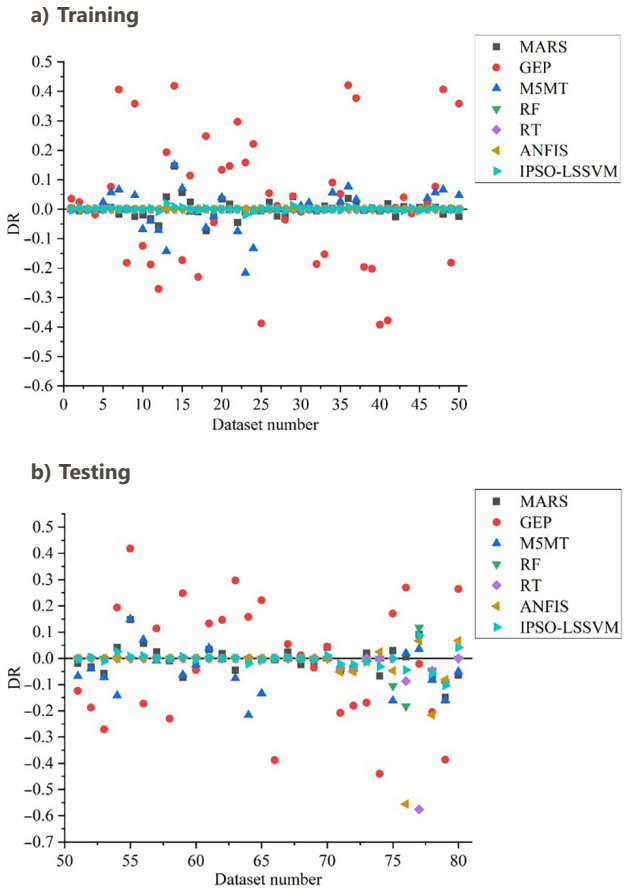

Figure 6. DR diagram of seven AI models

4.3. Development of stacked ensemble model

As can be seen from the above, out of the seven AI models, IPSO-LSSVM, RF, and RT models performed better, so we used these three AI models to form a stacked ensemble model for more robust comparisons. RF and RT models were used as single-based algorithms, and the IPSO-LSSVM model was employed as a meta-leaner in the stacked ensemble model. Five-fold cross-validation is used for the evaluation of the proposed models. The flowchart of the methodology is presented in Figure 7. The performance of the stacked ensemble model is shown in Table 7.

It can be seen from Table 7 that the R^2, RMSE, MAE, RRMSE and BIAS values of the stacked ensemble model are 0.995, 23.826 kN, 56 kN, 0.0417 and 0, respectively, for the training datasets and 0.9407, 120 kN, 184 kN, 0.1402 and 99.4834, respectively, for the test datasets. Compared with the RT, RF and IPSO-LSSVM models, the performance of the stacked ensemble model is not good, indicating that the stacked ensemble model is not suitable for the datasets used in this study.

4.4. Compared with models in literature

The superior performance of the IPSO-LSSVM model can be attributed to its powerful global search capability, especially when dealing with non-linear and high-dimensional data sets. The model adopts dynamic adjustment strategy to accelerate the convergence speed and reduce the risk of overfitting. This approach enhances the generalization power of the model, enabling it to adapt and perform well in a variety of prediction problems. Table 8 lists the comparison between the predicted results in this study and those in the literature. It can be seen from Table 8 that the R^2 of IPSO-LSSVM model is the highest, indicating that the IPSO-LSSVM model proposed in this study has certain advantages compared with other research models.

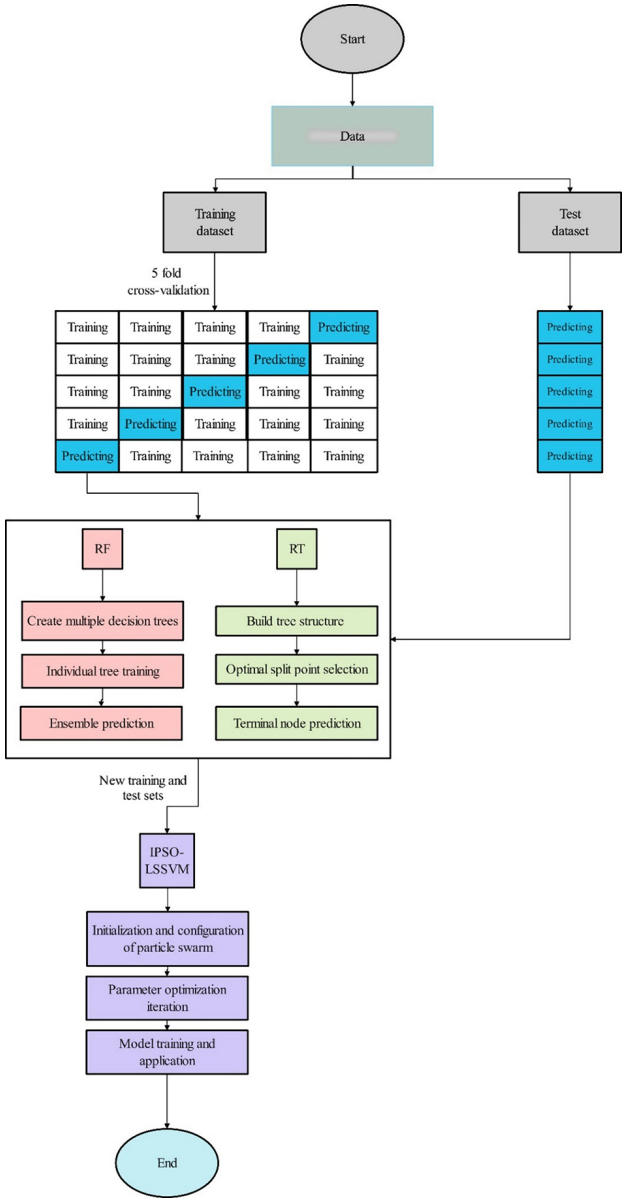

Figure 7. Flowchart for stacked ensemble model implementation

Table 7. Performance of the stacked ensemble model

Statistical indices	Training set	Test set
R^2	0.995	0.9407
MAE/kN	23.826	120
RMSE/kN	56	184
RRMSE	0.0417	0.1402
BIAS	0	99.4834

Table 8. Comparison with other models

Model	Reference	R^2
IPSO-LSSVM	This study	0.998
GA-ANN	Momeni et al. (2014)	0.990
Gaussian process (GP) regression	Pal and Deswal (2010)	0.950
ANN	Pham et al. (2020)	0.811
RF	Pham et al. (2020)	0.866
ANFIS	Yong et al. (2021)	0.910
GP	Yong et al. (2021)	0.971
SA-GP	Yong et al. (2021)	0.981
XGBoost	Amjad et al. (2022)	0.955
AdaBoost	Amjad et al. (2022)	0.950
DT	Amjad et al. (2022)	0.925
SVM	Amjad et al. (2022)	0.878
PSO-ANN	Ramesh et al. (2020)	0.993
DNN	Benbouras et al. (2021)	0.995
ELM	Benbouras et al. (2021)	0.869
Lasso	Benbouras et al. (2021)	0.925
PLS	Benbouras et al. (2021)	0.939
Kridge	Benbouras et al. (2021)	0.890
Ridge	Benbouras et al. (2021)	0.885
LS	Benbouras et al. (2021)	0.886
SVR	Benbouras et al. (2021)	0.868

4.5. Sensitivity analysis

To investigate the influence of the input parameters (i.e., D, L, $\bar{q}_{sik}$ and q_{pk}) on the prediction performance, four hybrid IPSO-LSSVM models were developed with different kinds of inputs. The statistical indices of the IPSO-LSSVM model on the training and test data sets are shown in Table 9.

As observed from Table 9, regardless of the training or testing datasets, the IPSO-LSSVM model with all the parameters has the highest R^2 and the lowest MAE, RMSE, RRMSE and BIAS, indicating that the IPSO-LSSVM with all the parameters has the best performance. For the test data sets, the R^2 of the IPSO-LSSVM model without D is 0.7601 lower than that of the IPSO-LSSVM with all parameters. The MAE, RMSE and RRMSE values of IPSO-LSSVM model without D are greater than those of the IPSO-LSSVM with all parameter, indicating that pile diameter D has the most significant influence on pile bearing capacity.

4.6. Performance of uncertainty, reliability and resilience analysis

In the process of using IPSO algorithm to optimize LSSVM, a number of RBF kernel parameters are generated. In this study, five typical kernel parameters (e.g., σ = 0.7448, 0.3189, 0.0174, 0.8538, 0.4482, respectively) are selected for comparative analysis in three aspects: uncertainty, reliability and resilience. The index U95 is used for uncertainty analysis. Specifically, U95 restricts the uncertainty of pile bearing capacity R_{uk} at a 95% confidence level. Obviously, the smaller the U95, the more accurate the pile bearing capacity R_{uk}. The other two indices, Reliability and Resilience, are used for reliability and resilience analysis respectively. The higher the value of these two indices, the more accurate the prediction result. A brief introduction to uncertainty, reliability and resilience analysis is given below (Saberi-Movahed et al., 2020).

Uncertainly analysis (Saberi-Movahed et al., 2020):

$$\mathrm{U95} = \left(\frac{1.96}{N}\right)\sqrt{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2 + \sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}, \tag{27}$$

where y_i and $\hat{y}_i$ are the ith measurement and the ith predicted values, respectively, $\bar{y}$ is the average of the measurement, N is the total number of samples.

Reliability analysis (Saberi-Movahed et al., 2020):

$$\mathrm{Reliability} = \left(\frac{100\%}{N}\right)\sum_{i=1}^{N} k_i, \tag{28}$$

where k_i is determined by the relative average error (RAE), which is defined as a vector whose ith component can be written as (Saberi-Movahed et al., 2020):

$$\mathrm{RAE}_i = \left|\frac{y_i - \hat{y}_i}{y_i}\right|. \tag{29}$$

If RAE_i is less than 0.05 (i.e., 5% error) set for this study, k_i is 1, otherwise it is 0.

Resilience analysis (Saberi-Movahed et al., 2020):

$$\mathrm{Resiliency} = \left(\frac{\sum_{i=1}^{N-1} r_i}{N - \sum_{i=1}^{N} k_i}\right) \times 100\%, \tag{30}$$

where r_i is the probable number of times the prediction model can recover from an inaccurate estimate to an accurate estimate in the ith data sample.

Table 9. Single parameter sensitivity analysis of the IPSO-LSSVM model

Model	Training					Testing				
	R^2	MAE/kN	RMSE/kN	RRMSE	BIAS	R^2	MAE/kN	RMSE/kN	RRMSE	BIAS
With all parameters	1	4.27	6.164	0.005	0	0.9977	22	36.03	0.0275	–11
Without D	0.9996	28	32.42	0.0241	0	0.7601	195	370	0.282	–72.96
Without L	0.9997	14	17.57	0.0131	0	0.9406	83	184	0.14	–38.37
Without $\bar{q}_{sik}$	0.9987	27	39.17	0.0291	0	0.7709	177	357	0.272	–75.02
Without q_{pk}	0.9993	21	26.78	0.0199	0	0.8476	140	295	0.225	–53.71

Table 10. Uncertainty, reliability and resiliency results of the proposed RBF network

Model	Training			Testing		
	U95	Reliability (%)	Resilience (%)	U95	Reliability (%)	Resilience (%)
IPSO-LSSVM (σ = 0.7448)	219.094	100	100	271.084	90	98
IPSO-LSSVM (σ = 0.3189)	219.138	96	99	271.277	77	85
IPSO-LSSVM (σ = 0.0174)	219.087	100	100	289.97	70	70
IPSO-LSSVM (σ = 0.8538)	219.157	98	98	271.258	86.7	80
IPSO-LSSVM (σ = 0.4482)	219.089	100	100	271.105	90	80

The uncertainty, reliability and resiliency results of the proposed RBF network are shown in Table 10.

It can be seen from Table 10 that the U95, reliability and resilience values of the IPSO-LSSVM (σ = 0.7448) model are 219.094, 100% and 100%, respectively, for the training datasets, and 271.084, 90% and 98%, respectively, for the testing datasets. In the testing stage, compared to other IPSO-LSSVM models, the U95 of the IPSO-LSSVM (σ = 0.7448) model is the smallest, while reliability and resilience are the largest, indicating that the IPSO-LSSVM model (σ = 0.7448) has the best predictive performance. Meanwhile, it can be found that the values of different kernel parameters will affect the prediction performance of the model, so it is necessary to choose the optimal kernel parameters.

5. Conclusions

This study presents a hybrid IPSO-LSSVM model for the prediction of bearing capacity of piles. The improved PSO algorithm was used to optimize the LSSVM hyperparameters. The performance of the IPSO-LSSVM model was compared with seven AI models, ANFIS, M5MT, MARS, GEP, RF, RT and a stacked ensemble model. Six statistical indices (e.g., R^2, MAE, RMSE, RRMSE, BIAS and DR) were used to evaluate the performance of the models. The R^2, MAE, RMSE, RRMSE and BIAS values of the IPSO-LSSVM model were 1, 4.27 kN, 6.164 kN, 0.005 and 0, respectively, for the training datasets and 0.9977, 22 kN, 36.03 kN, 0.0275 and –11, respectively, for the testing datasets. Compared with the ANFIS, MARS, GEP, M5MT, RF, RT and the stacked ensemble models, the proposed IPSO-LSSVM model shows high accuracy and robustness on the test datasets.

Compared with the literature results, the IPSO-LSSVM model has higher prediction accuracy, and the R^2 of the test datasets is 0.998.

The sensitivity analysis results show that the MAE, RMSE and RRMSE values of IPSO-LSSVM model without *D* are greater than those of the IPSO-LSSVM with all parameter, indicating that pile diameter *D* has the most significant influence on pile bearing capacity.

The uncertainty, reliability and resiliency results show that the U95, Reliability and Resilience values of the IPSO-LSSVM (σ = 0.7448) model are 219.094, 100% and 100%, respectively, for the training datasets, and 271.084, 90% and 98%, respectively, for the testing datasets. In the testing stage, compared to other IPSO-LSSVM models, the U95 of the IPSO-LSSVM (σ = 0.7448) model is the smallest, while Reliability and Resilience are the largest, indicating that the IPSO-LSSVM model (σ = 0.7448) has the best predictive performance.

One limitation of this study is that the proposed model is only applicable to the current range of data and parameters, and cannot guarantee high prediction accuracy when the data exceeds the predetermined range. In addition, the values of different kernel parameters will affect the prediction performance of the model, so it is necessary to choose the optimal kernel parameters. Therefore, the accuracy of the proposed IPSO-LSSVM model still needs to be improved, and more data samples are needed for calibration.

Author contribution

Li Tao: Conceptualization, Methodology, Writing-original draft. Xinhua Xue: Conceptualization, Methodology, Writing-reviewing, Editing, Supervision.

Funding

No funding is associated with this article.

Data availability

Data will be made available on reasonable request.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

Ahmadi, M. H., Baghban, A., Ghzavini, M., Hadipoor, M., Ghasempour, R., & Nazemzadegan, M. R. (2020). An insight into the prediction of TiO_2/water nanofluid viscosity through intelligence schemes. *Journal of Thermal Analysis and Calorimetry*, *139*, 2381–2394. https://doi.org/10.1007/s10973-019-08636-4

Alkroosh, I., & Nikraz, H. (2012). Predicting axial capacity of driven piles in cohesive soils using intelligent computing. *Engineering Applications of Artificial Intelligence*, *25*(3), 618–627. https://doi.org/10.1016/j.engappai.2011.08.009

Alkroosh, I., & Nikraz, H. (2014). Predicting pile dynamic capacity via application of an evolutionary algorithm. *Soils and Foundations*, *54*(2), 233–242. https://doi.org/10.1016/j.sandf.2014.02.013

Amjad, M., Ahmad, I., Ahmad, M., Wróblewski, P., Kaminski, P., & Amjad, U. (2022). Prediction of pile bearing capacity using XGBoost algorithm: modeling and performance evaluation. *Applied Sciences*, *12*(4), Article 2126. https://doi.org/10.3390/app12042126

Armaghani, D. J., Shoib, R. S. N. S. B. R., Faizi, K., & Rashid, A. S. A. (2017). Developing a hybrid PSO-ANN model for estimating the ultimate bearing capacity of rock-socketed piles. *Neural Computing and Applications*, *28*, 391–405. https://doi.org/10.1007/s00521-015-2072-z

Baziar, M. H., Kashkooli, A., & Azizkandi, A. S. (2012). Prediction of pile shaft resistance using cone penetration tests (CPTs). *Computers and Geotechnics*, *45*, 74–82. https://doi.org/10.1016/j.compgeo.2012.04.005

Baghban, A., & Khoshkharam, A. (2016). Application of LSSVM strategy to estimate asphaltene precipitation during different production processes. *Petroleum Science and Technology*, *34*(22), 1855–1860. https://doi.org/10.1080/10916466.2016.1237966

Baghban, A., Kashiwao, T., Bahadori, M., Ahmad, Z., & Bahadori, A. (2016a). Estimation of natural gases water content using adaptive neuro-fuzzy inference system. *Petroleum Science and Technology*, *34*(10), 891–897. https://doi.org/10.1080/10916466.2016.1176039

Baghban, A., Bahadori M., Ahmad, Z., Kashiwao, T., & Bahadori, A. (2016b). Modelling of true vapor pressure of petroleum products using ANFIS algorithm. *Petroleum Science and Technology*, *34*(10), 933–939. https://doi.org/10.1080/10916466.2016.1170843

Baghban, A., Abbasi, P., & Rostami, P. (2016c). Modeling of viscosity for mixtures of Athabasca bitumen and heavy n-alkane with LSSVM algorithm. *Petroleum Science and Technology*, *34*(20), 1698–1704. https://doi.org/10.1080/10916466.2016.1219748

Bahadori, A., Baghban, A., Bahadori, M., Lee, M., Ahmad, Z., Zare, M., & Abdollahi, E. (2016). Computational intelligent strategies to predict energy conservation benefits in excess air controlled gas-fired systems. *Applied Thermal Engineering*, *102*, 432–446. https://doi.org/10.1016/j.applthemaleng.2016.04.005

Bemani, A., Baghban, A., Mosavi, A., & Shahab, S. (2020a). Estimating CO_2-Brine diffusivity using hybrid models of ANFIS and evolutionary algorithms. *Engineering Applications of Computational Fluid Mechanics*, *14*(1), 818–834. https://doi.org/10.1080/19942060.2020.1774422

Bemani, A., Baghban, A., Mohammadi, & Amir H. (2020b). An insight into the modeling of sulfur content of sour gases in supercritical region. *Journal of Petroleum Science and Engineering*, *184*, Article 106459. https://doi.org/10.1016/j.petrol.2019.106459

Bemani, A., Baghban, A., Shamshirband, S., Mosavi, A., Csiba, P., & Varkonyi-Koczy, A. R. (2020c). Applying ANN, ANFIS, and LSSVM models for estimation of acid solvent solubility in supercritical CO_2. *Computers*, *Materials* & *Continua*, *63*(3), 1175–1204. https://doi.org/10.32604/cmc.2020.07723

Benbouras, M. A., Petrişor. A.-I., Zedira, H., Ghelani, L., & Lefilef, L. (2021). Forecasting the bearing capacity of the driven piles using advanced machine-learning techniques. *Applied Sciences*, *11*(22), Article 10908. https://doi.org/10.3390/app112210908

Breiman, L., Friedman, J. H., Olshen, B. A., & Stone, C. (1984). Classification and regression trees. *Biometrics*, *40*, Article 874. https://doi.org/10.2307/2530946

Chiu, S. L. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent* & *Fuzzy Systems: Applications in Engineering and Technology*, *2*(3), 267–278. https://doi.org/10.3233/IFS-1994-2306

Das, S. K., & Basudhar, P. K. (2006). Undrained lateral load capacity of piles in clay using artificial neural network. *Computers and Geotechnics*, *33*(8), 454–463. https://doi.org/10.1016/j.compgeo.2006.08.006

Daneshvar, D., & Behnood, A. (2020). Estimation of the dynamic modulus of asphalt concretes using random forests algorithm. *International Journal of Pavement Engineering*, *23*(2), 250–260. https://doi.org/10.1080/10298436.2020.1741587

Ferreira, C. (2001). Gene expression programming: a new adaptive algorithm for solving problems. *Complex Systems*, *13*(2), 87–129. https://doi.org/10.48550/arXiv.cs/0102027

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, *19*(1), 1–67. https://doi.org/10.1214/aos/1176347963

Haratipour, P., Baghban, A., Mohammadi, A. H., Nazhad, S. H. H., & Bahadori, A. (2017). On the estimation of viscosities and densities of CO_2-loaded MDEA, MDEA+AMP, MDEA+DIPA, MDEA+MEA, and MDEA+DEA aqueous solutions. *Journal of Molecular Liquids*, *242*, 146–159. https://doi.org/10.1016/j.moliq.2017.06.123

Homaei, F., & Najafzadeh, M. (2020). A reliability-based probabilistic evaluation of the wave-induced scour depth around marine structure piles. *Ocean Engineering*, *196*, Article 106818. https://doi.org/10.1016/j.oceaneng.2019.106818

Homaei, F., & Najafzadeh, M. (2022). Failure analysis of scouring at pile groups exposed to steady-state flow: On the assessment of reliability-based probabilistic methodology. *Ocean Engineering*, *266*(Part 3), Article 112707. https://doi.org/10.1016/j.oceaneng.2022.112707

Kalinli, A., Acar, M. C., & Gunduz, Z. (2011). New approaches to determine the ultimate bearing capacity of shallow foundations based on artificial nerual networks and ant colony optimization. *Engineering Geology*, *117*(1–2), 29–38. https://doi.org/10.1016/j.enggeo.2010.10.002

Kardani, M. N., Baghban, A., Sasanipour, J., Mohammadi, Amir, H., & Habibzadeh, S. (2018). Group contribution methods for estimating CO_2 absorption capacities of imidazolium and ammonium-based polyionic liquids. *Journal of Cleaner Production*, *203*, 601–618. https://doi.org/10.1016/j.jclepro.2018.08.127

Kennedy, J., & Eberhart, R. C. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 – International Conference on Neural Networks* (Vol. 4, pp. 1942–1948), Perth, Australia. IEEE Service Center. https://doi.org/10.1109/ICNN.1995.488968

Li, L., Li, J. P., Sun, D. A., & Zhang, L. X. (2017). Time-dependent bearing capacity of a jacked pile: An analytical approach based on effective stress method. *Ocean Engineering*, *143*, 177–185. https://doi.org/10.1016/j.oceaneng.2017.08.010

Lin, H. M., Chang, S. K., Wu, S. K., & Juang, C. H. (2009). Neural network-based model for assessing failure potential of highway slopes in the Alishan, Taiwan Area: Pre- and post earthquake investigation. *Engineering Geology*, *104*(3–4), 280–289. https://doi.org/10.1016/j.enggeo.2008.11.007

Luo, R. P., Yang, M., & Li, W. C. (2018). Normalized settlement of piled raft in homogeneous clay. *Computers and Geotechnics*, *103*, 165–178. https://doi.org/10.1016/j.compgeo.2018.07.023

Mercer, J. (1909). Functions of positive and negative type, and their connection with the theory of integral equations. *Proceedings of the Royal Society A*, *209*, 415–446. https://doi.org/10.1098/rspa.1909.0075

Momeni, E., Nazir, R., Armaghani, D. J., & Maizir, H. (2014). Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. *Measurement*, *57*, 122–131. https://doi.org/10.1016/j.measurement.2014.08.007

Murlidhar, B. R., Sinha, R. K., Mohamad, E. T., Sonkar, R., & Khorami, M. (2020). The effects of particle swarm optimisation and genetic algorithm on ANN results in predicting pile bearing capacity. *International Journal of Hydromechatronics*, *3*(1), 69–87. https://doi.org/10.1504/IJHM.2020.105484

Najafzadeh, M. (2015). Neuro-fuzzy GMDH systems based evolutionary algorithms to predict scour pile groups in clear water conditions. *Ocean Engineering*, *99*, 85–94. https://doi.org/10.1016/j.oceaneng.2015.01.014

Najafzadeh, M., & Azamathulla, H. M. (2013a). Group method of data handling to predict scour depth around bridge piers. *Neural Computing and Applications*, *23*, 2107–2112. https://doi.org/10.1007/s00521-012-1160-6

Najafzadeh, M., & Azamathulla, H. M. (2013b). Neuro-fuzzy GMDH to predict the scour pile groups due to waves. *Journal of Computing in Civil Engineering*, *29*(5), Article 04014068. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000376

Najafzadeh, M., & Barani, G.-A. (2011). Comparison of group method of data handling based genetic programming and back propagation system to predict scour depth around bridge piers. *Scientia Iranica*, *18*(6), 1207–1213. https://doi.org/10.1016/j.scient.2011.11.017

Najafzadeh, M., & Oliveto, G. (2021). More reliable predictions of clear-water scour depth at pile groups by robust artificial intelligence techniques while preserving physical consistency. *Soft Computing*, *25*, 5723–5746. https://doi.org/10.1007/s00500-020-05567-3

Najafzadeh, M., & Mahmoudi-Rad, M. (2024). New empirical equations to assess energy efficiency of flow-dissipating vortex dropshaft. *Engineering Applications of Artificial Intelligence*, *131*, Article 107759. https://doi.org/10.1016/j.engappai.2023.107759

Najafzadeh, M., Barani, G.-A., & Azamathulla, H. M. (2013). GMDH to predict scour depth around a pier in cohesive soils. *Applied Ocean Research*, *40*, 35–41. https://doi.org/10.1016/j.apor.2012.12.004

Najafzadeh, M., Etemad-Shahidi, A., & Lim, S. Y. (2016). Scour prediction in long contractions using ANFIS and SVM. *Ocean Engineering*, *111*, 128–135. https://doi.org/10.1016/j.oceaneng.2015.10.053

Nejad, F. P., & Jaksa, M. B. (2017). Load-settlement behavior modeling of single piles using artificial neural networks and CPT data. *Computers and Geotechnics*, *89*, 9–21. https://doi.org/10.1016/j.compgeo.2017.04.003

Pal, M., & Deswal, S. (2010). Modelling pile capacity using Gaussian process regression. *Computers and Geotechnics*, *37*, 942–947. https://doi.org/10.1016/j.compgeo.2010.07.012

Pham, T. A., Ly, H.-B., Tran, V. Q., & Giap, L. V. (2020). Prediction of Pile Axial bearing capacity using artificial neural network and random forest. *Applied Sciences*, *10*(5), Article 1871. https://doi.org/10.3390/app10051871

Quinlan, J. R. (1992). Learning with continuous classes. In *Proceedings of AI'92* (pp. 343–348). Singapore.

Ramesh, M. B., Kumar, S. R., Tonnizam, M. E., Rajesh, S., & Majid, K. (2020). The effects of particle swarm optimisation and genetic algorithm on ANN results in predicting pile bearing capacity. *International Journal of Hydromechatronics*, *3*(1), 69–87. https://doi.org/10.1504/IJHM.2020.105484

Rezazadeh, S., & Eslami, A. (2017). Empirical methods for determining shaft bearing capacity of semi-deep foundations socketed in rocks. *Journal of Rock Mechanics and Geotechnical Engineering*, *9*(6), 1140–1151. https://doi.org/10.1016/j.jrmge.2017.06.003

Saberi-Movahed, F., Najafzadeh, M., & Mehrpooya, A. (2020). Receiving more accurate prediction for longitudinal dispersion coefficients in water pipelines: Training group method of data handling using extreme learning machine conceptions. *Water Resources Management*, *34*, 529–561. https://doi.org/10.1007/s11269-019-02463-w

Salgado, R., Zhang, Y. B., Abou-Jaoude, G., Loukidis, D., & Bisht, V. (2017). Pile driving formulas based on pile wave equation analyses. *Computers and Geotechnics*, *81*, 307–321. https://doi.org/10.1016/j.compgeo.2016.09.004

Sheil, B. B., & McCabe, B. A. (2016). An analytical approach for the prediction of single pile and pile group behavior in clay. *Computers and Geotechnics*, *75*, 145–158. https://doi.org/10.1016/j.compgeo.2016.02.001

Suykens, J. A. K, Vandewalle, J., & De Moor, B. (2001). Optimal control by least squares support vector machines. *Neural Networks*, *14*(1), 23–35. https://doi.org/10.1016/S0893-6080(00)00077-0

Teh, C. I., Wong, K. S., Goh, A. T. C., & Jaritngam, S. (1997). Prediction of pile capacity using neural networks. *Journal of Computing in Civil Engineering*, *11*(2), 129–138. https://doi.org/10.1061/(ASCE)0887-3801(1997)11:2(129)

Wang, X. F., Zeng, X. W., & Li, J. L. (2018). Assessment of bearing capacity of axially loaded monopiles based on centrifuge tests. *Ocean Engineering*, *167*, 357–368. https://doi.org/10.1016/j.oceaneng.2018.08.063

Yong, W. X., Zhou, J., Armaghani, D. J., Tahir, M. M., Tarinejad, R., Pham, B. T., & Huynh, V. V. (2021). A new hybrid simulated annealing-based genetic programming technique to predict the ultimate bearing capacity of piles. *Engineering with Computers*, *37*, 2111–2127. https://doi.org/10.1007/s00366-019-00932-9

Zheng, R. Y., Wu, S., & Wang, N. J. (2006). Predicting ultimate bearing capacity of single pile using ANFIS and reliability analysis. *Industrial Construction*, *36*(6), 70–76 (in Chinese).

APPENDIX

Table A1. Experimental datasets

Datasets	No.	D/mm	L/m	$\overline{q}_{sik}$/kPa	q_{pk}/kPa	R_{uk}/kN
Training	1	500	47	20.7	1400	1900
	2	500	45	18.5	1400	1650
	3	500	47	20.8	2000	2000
	4	500	45	19.2	2000	1800
	5	400	39	9.345	3000	800
	6	500	31	8.5	600	500
	7	400	28	13	800	500
	8	550	51	19.9	3000	2500
	9	400	24	12.8	850	450
	10	500	30	12.9	850	900
	11	500	35	11.7	1900	1050
	12	500	29	12.8	1800	1050
	13	400	27	10.4	1300	500
	14	400	30	12	1300	500
	15	500	26	13.3	1800	900
	16	500	46	13.3	800	1100
	17	600	57	18.5	2300	2700
	18	400	38	11.4	1300	700
	19	500	40	11.6	1300	1000
	20	500	47	12.6	1000	1050
	21	500	49	13.5	800	1150
	22	400	34	11.5	1000	600
	23	400	27	10.3	1200	510
	24	400	31	10.5	1200	550
	25	600	38.5	21	3200	2500
	26	500	47	21.1	1400	1900
	27	500	45	18.3	1300	1650
	28	500	47	20.1	2100	2000
	29	500	48	19.2	2000	1800
	30	400	39	9.2	3000	800
	31	400	40	9.1	3000	800
	32	550	51	19.8	3000	2500
	33	550	52	20.1	3000	2500
	34	500	32	8.4	550	500
	35	500	30	8.6	550	500
	36	400	26	12.7	800	450
	37	400	25	12.6	800	450
	38	600	59	18.5	2400	2700
	39	600	58	18.7	2400	2700
	40	600	38	21.1	3400	2500
	41	600	39	20.9	3400	2500
	42	500	46	20.5	1400	1900
	43	500	46	18.4	1400	1650
	44	500	46	21	2000	2000
	45	500	46	19.1	2000	1800
	46	400	40	9.3	3000	800
	47	500	31	8.5	600	500
	48	400	28	13	800	500

End of Table A1

Datasets	No.	D/mm	L/m	$\overline{q}_{sik}$/kPa	q_{pk}/kPa	R_{uk}/kN
Training	49	550	51	19.9	3000	2500
	50	400	24	12.8	850	450
	51	500	30	12.9	850	900
	52	500	35	11.7	1900	1050
	53	500	29	12.8	1800	1050
	54	400	27	10.4	1300	500
	55	400	30	12	1300	500
	56	500	26	13.3	1800	900
	57	500	46	13.3	800	1100
	58	600	57	18.5	2300	2700
	59	400	38	11.4	1300	700
	60	500	40	11.6	1300	1000
Testing	61	500	47	12.6	1000	1050
	62	500	49	13.5	800	1150
	63	400	34	11.5	1000	600
	64	400	27	10.3	1200	510
	65	400	31	10.5	1200	550
	66	600	38.5	21	3200	2500
	67	500	47	21.1	1400	1900
	68	500	45	18.3	1300	1650
	69	500	47	20.1	2100	2000
	70	500	48	19.2	2000	1800
	71	550	52	19.8	2900	2600
	72	550	53	20	2900	2600
	73	600	59	19	2500	2700
	74	600	36	21.1	3400	2500
	75	400	26	10.5	1300	500
	76	400	31	12	1400	600
	77	500	30	12.7	850	800
	78	500	36	11.8	1800	1100
	79	500	28	12.8	1900	1150
	80	400	39	11.3	1300	700